

CACHE LINE OWNERSHIP TRANSFER IN MULTI-PROCESSOR COMPUTER SYSTEMS

BACKGROUND

[0001] In shared memory, multi-processor computer systems, cache miss latency has a significant effect on system performance. (In the context of the invention, “processor” includes, but is not limited to, central processing units (CPUs) and I/O processing agents.) As those skilled in the art will understand, a “cache miss” occurs when a processor checks its cache for data and discovers that the desired data is not in the cache. A “cache miss” is the opposite of a “cache hit,” which occurs when the requested information is in the cache. If a cache miss occurs, the processor must request the desired data, referred to as a “cache line,” from the computer system’s memory subsystem. The time it takes a processor to check its cache, discover that the data is not in the cache, request the desired data from the memory subsystem, and receive the data from the memory subsystem, is time during which the processor is idle, and is referred to as cache miss latency.

[0002] In a large system, cache miss latency can be extremely large, particularly where a processor requests ownership of a cache line owned by a different processor located at a remote cell. A cell is a sub-module of the system and typically has a number of system resources, such as central processing units (CPUs), central agent controllers, input/output (I/O) processing units, and memory. Cells can be configured as a single shared memory domain or as part of multiple cells grouped together to form a shared memory domain. Several steps are involved in transferring ownership of a cache line between processors, and each step increases cache miss latency.

SUMMARY

[0003] The invention aims to reduce cache miss latency by reducing the number of steps in transferring ownership of a cache line, thus reducing latency. In one aspect, the invention encompasses a method of transferring ownership of a cache line between processors in a shared memory multi-processor computer system. The method comprises sending a request for ownership of a cache line from a first processor to a memory unit. The memory unit receives the request and determines which one of a plurality of processors other than the first processor has ownership of the requested cache line. The memory sends a recall for ownership to that other processor. The cache line data with ownership is sent from the

other processor to the first processor in response to the recall. A response may be sent from the first processor to the memory unit to confirm receipt of the ownership of the requested cache line by the first processor.

[0004] Optionally, an additional response may be sent from the other processor to the memory unit to confirm that the other processor has sent the ownership of the requested cache line to the first processor. A copy of the requested cache line data may, but need not always, be sent to the memory unit as part of this additional response.

[0005] The invention encompasses both cell-based and non cell-based computer systems. For cell-based systems, the invention encompasses both single cell shared memory systems and multiple cell systems forming a single shared memory domain. The processors and memory unit may reside on one, two, or three distinct cells, in any grouping.

BRIEF DESCRIPTION OF THE DRAWINGS


[0006] Figure 1 is a simplified flow diagram showing some of the steps in requesting ownership of a cache line between processors in a multi-processor directory-based cache coherency system.

[0007] Figure 2 is a simplified flow diagram illustrating a method according to one embodiment of the present invention.

[0008] The arrows in Figures 1 and 2 represent transactions sent between the processors and memory. The invention encompasses all implementations by which those transactions are sent, including packets, signals, busses, messages, and the like. The term "transaction" is used herein to cover all such implementations.

DETAILED DESCRIPTION

[0009] Figure 1 illustrates in greatly simplified form a shared memory multiprocessor computer system that uses a directory-based cache coherency scheme. The simplified system is divided into three cells, arbitrarily designated Cell 0, Cell 1, and Cell 2. A first processor, arbitrarily designated Processor A, is associated with Cell 0. Another processor, arbitrarily designated Processor B, is associated with Cell 2. Both Processor A



and Processor B have their own cache memory. A memory unit is associated with Cell 1 and is shared by both Processor A and Processor B.

[0010] If Processor A requires ownership of a cache line owned by the memory unit, a request transaction, which contains the memory address for the requested cache line stored in the memory unit, is sent from Processor A to the memory unit that owns the cache line, as represented by arrow 10. The memory unit receives the request transaction and determines from the DRAM tag for the memory address of the requested cache line that Processor B, associated with Cell 2, has ownership of the requested cache line. The memory unit then recalls the requested cache line out of Processor B's cache by sending a recall transaction, represented by arrow 20, to Processor B. In response, Processor B returns the cache line data and ownership of the requested cache line to the memory unit by sending a response transaction, as represented by arrow 30. Then, the memory unit transfers the cache line data and ownership of the requested cache line to Processor A by sending a data transaction, as represented by arrow 40.

[0011] A disadvantage of this method of operation is that, while ownership of the cache line is being requested by Processor A and transferred from Processor B to Processor A, the process running on Processor A is stalled until ownership of the requested cache line and data are sent to Processor A. The path represented by arrows 10, 20, 30, and 40 is referred to herein as the "latency critical path," and the time the process running on Processor A is stalled is referred to as "cache miss latency."

[0012] A method for reducing cache miss latency according to the invention is illustrated in greatly simplified form in Figure 2. According to the embodiment illustrated in Figure 2, one of the steps in the latency critical path is removed. As illustrated in Figure 2, if Processor A requires ownership of a cache line owned by the memory unit, a request transaction, which contains the memory address, is sent from Processor A to the memory unit, as represented by arrow 10. The memory unit receives the request transaction and determines from the DRAM tag for the memory address of the requested cache line that Processor B, associated with Cell 2, has ownership of the requested cache line. The memory unit then sends a recall transaction, represented by arrow 20, to Processor B to recall the requested cache line out of Processor B's cache.

[0013] In response to the recall transaction, Processor B sends the cache line data and ownership of the requested cache line to the requesting processor, Processor A, by sending a data transaction, as represented by arrow 30a.

[0014] A copy of the cache line is sent to the memory unit to update the cache line held in the memory unit by sending a response transaction, as represented by arrow 30b. However, that is not necessary in all cases. In some cases, Processor B need not send a copy of the cache line to the memory unit. When Processor A requests ownership of a cache line, it can accompany that request with an indication that it either will, if requested, guarantee to provide that cache line to a different processor on a subsequent request, or will not guarantee to provide the cache line. If in the request for ownership of the cache line Processor A guarantees to provide the cache line in response to a subsequent request, the response from Processor B to the memory unit, represented by 30b, is sent to the memory unit without a copy of the cache line data. When used, this approach greatly reduces system bandwidth. However, this approach can only be used when Processor A guarantees in the initial request that the cache line data will be provided upon receipt of a subsequent request. If in the initial request Processor A does not guarantee it will provide the cache line on a subsequent request, the response indicated by 30b is sent to the memory unit with a copy of the cache line data.

[0015] To complete the coherency flow, a response transaction, represented by arrow 50, may be sent from Processor A to the memory unit. The transaction informs the memory that the cache line has been received by the original requesting processor, Processor A. In response, the memory unit updates the DRAM tag to indicate that ownership of the cache line has been transferred to Processor A.

[0016] Additionally, the response from Processor B to the memory unit, represented by 30b, can be omitted entirely when Processor A guarantees in the initial ownership request that the cache line data will be provided upon receipt of a subsequent request. In that event, the response transaction from Processor A, which informs the memory that the cache line has been received from Processor B, also necessarily informs the memory that Processor B has sent the cache line to Processor A. Thus, in that event the response represented by arrow 30b is not required.

[0017] The transaction flow depicted in Figure 2 can also be carried out when Processor A, Processor B, and the memory unit reside on one or two cells in any manner.

The transaction flow can also be carried out in a non cell-based system architecture. In addition, the invention is not limited to any particular number of processors or memory units.

[0018] The invention reduces cache miss latency in a multiprocessor system. The reduced idle time for stalled processes waiting for the requested data contained in a cache line allows applications and benchmarks to run significantly faster.

[0019] Although the present invention and its advantages have been described in detail, it should be understood that various changes, substitutions, and alterations can be made without departing from the spirit and scope of the invention as defined by the appended claims. Moreover, the scope of the present application is not intended to be limited to the particular embodiments of invention described in the specification. As one of ordinary skill in the art will readily appreciate from the foregoing description, processes, machines, articles of manufacture, compositions of matter, means, methods, or steps presently existing or later to be developed that perform substantially the same function or achieve substantially the same result as the corresponding embodiments described herein may be utilized to implement and carry out the present invention. Accordingly, the appended claims are intended to include within their scope such processes, machines, articles of manufacture, compositions of matter, means, methods, or steps.

[0020] The foregoing describes the invention in terms of embodiments foreseen by the inventors for which an enabling description was available, notwithstanding that insubstantial modifications of the invention, not presently foreseen, may nonetheless represent equivalents thereto.